



**University of
Zurich^{UZH}**

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2012

Quantifying the importance of local niche-based and stochastic processes to tropical tree community assembly

Shipley, Bill ; Paine, C E Timothy ; Baraloto, Christopher

Abstract: Although niche-based and stochastic processes, including dispersal limitation and demographic stochasticity, can each contribute to community assembly, it is difficult to quantify the relative importance of each process in natural vegetation. Here, we extend Shipley's maxent model (Community Assembly by Trait Selection, CATS) for the prediction of relative abundances to incorporate both trait-based filtering and dispersal limitation from the larger landscape and develop a statistical decomposition of the proportions of the total information content of relative abundances in local communities that are attributable to trait-based filtering, dispersal limitation, and demographic stochasticity. We apply the method to tree communities in a mature, species-rich, tropical forest in French Guiana at 1-, 0.25- and 0.04-ha scales. Trait data consisted of species' means of 17 functional traits measured over both the entire meta-community and separately in each of nine 1-ha plots. Trait means calculated separately for each site always gave better predictions. There was clear evidence of trait-based filtering at all spatial scales. Trait-based filtering was the most important process at the 1-ha scale (34%), whereas demographic stochasticity was the most important at smaller scales (37–53%). Dispersal limitation from the meta-community was less important and approximately constant across scales (9%), and there was also an unresolved association between site-specific traits and meta-community relative abundances. Our method allows one to quantify the relative importance of local niche-based and meta-community processes and demographic stochasticity during community assembly across spatial and temporal scales.

DOI: <https://doi.org/10.1890/11-0944.1>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-78232>

Journal Article

Published Version

Originally published at:

Shipley, Bill; Paine, C E Timothy; Baraloto, Christopher (2012). Quantifying the importance of local niche-based and stochastic processes to tropical tree community assembly. *Ecology*, 93(4):760-769.

DOI: <https://doi.org/10.1890/11-0944.1>



ECOLOGICAL SOCIETY OF AMERICA

Ecology/Ecological Monographs/Ecological Applications

PREPRINT

This preprint is a PDF of a manuscript that has been accepted for publication in an ESA journal. It is the final version that was uploaded and approved by the author(s). While the paper has been through the usual rigorous peer review process of ESA journals, it has not been copy-edited, nor have the graphics and tables been modified for final publication. Also note that the paper may refer to online Appendices and/or Supplements that are not yet available. We have posted this preliminary version of the manuscript online in the interest of making the scientific findings available for distribution and citation as quickly as possible following acceptance. However, readers should be aware that the final, published version will look different from this version and may also have some differences in content.

The doi for this manuscript and the correct format for citing the paper are given at the top of the online (html) abstract.

Once the final published version of this paper is posted online, it will replace the preliminary version at the specified doi.

Quantifying the importance of local niche-based and stochastic processes to
tropical tree community assembly

Bill Shipley¹, C. E. Timothy Paine² and Christopher Baraloto³

¹Département de biologie, Université de Sherbrooke, Sherbrooke (Qc) Canada J1K 2R1,
Bill.Shipley@USherbrooke.ca

²Institut für Evolutionsbiologie und Umweltwissenschaften, Universität Zürich,
Winterthurerstrasse 190, 8057 Zürich, Switzerland
timothy.paine@ieu.uzh.ch

³INRA, UMR Écologie des Forêts de Guyane, 97387 Kourou Cedex, French Guiana
chris.baraloto@ecofog.gf

pages of text: 20

pages total: 30

words: 6540

tables: 3

figures: 1

Supplements: 1

Running head: niche and neutral community assembly

26 ABSTRACT

27 Although niche-based and stochastic processes, including dispersal limitation and
 28 demographic stochasticity, can each contribute to community assembly, it is difficult to
 29 quantify the relative importance of each process in natural vegetation. Here, we extend
 30 Shipley's maxent model (Community Assembly by Trait Selection, CATS) for the
 31 prediction of relative abundances to incorporate both trait-based filtering and dispersal
 32 limitation from the larger landscape and develop a statistical decomposition of the
 33 proportions of the total information content of relative abundances in local communities
 34 that are attributable to trait-based filtering, dispersal limitation and demographic
 35 stochasticity. We apply the method to tree communities in a mature, species-rich,
 36 tropical forest in French Guiana at 1, 0.25 and 0.04 ha scales. Trait data consisted of
 37 species' means of 17 functional traits measured over both the entire meta-community and
 38 separately in each of nine 1-ha plots. Trait means calculated separately for each site
 39 always gave better predictions. There was clear evidence of trait-based filtering at all
 40 spatial scales. Trait-based filtering was the most important process at the 1-ha scale
 41 (34%), whereas demographic stochasticity was the most important at smaller scales (37 -
 42 53%). Dispersal limitation from the meta-community was less important and
 43 approximately constant across scales (~9%), and there was also an unresolved association
 44 between site-specific traits and meta-community relative abundances. Our method
 45 allows one to quantify the relative importance of local niche-based and meta-community
 46 processes and demographic stochasticity during community assembly across spatial and
 47 temporal scales.

48 Key words: CATS, demographic stochasticity, dispersal limitation, environmental
49 filtering, French Guiana, functional traits, maxent, neutral assembly, tropical forests.

50

esa
preprint

51 INTRODUCTION

52 Neutral models (Bell 2000, Hubbell 2001) assume equal per capita demographic
 53 properties of all species. They thus presume that community assembly does not include
 54 niche-based processes, but arises via dispersal limitation from the meta-community and
 55 local demographic stochasticity. Deterministic models of community assembly (Lotka
 56 1925, Volterra 1926, 1931, MacArthur 1972, Tilman 1982, 1988) represent the other
 57 extreme, in which niche-based factors dominate. Between these extremes lies a
 58 continuum of possibilities in which community assembly is influenced to varying degrees
 59 by local niche-based, stochastic, and landscape processes (Gravel et al. 2006). It has long
 60 been recognized (Tansley 1920) that the non-random processes of competition and
 61 abiotic environmental filtering interact with the unpredictable processes of seed arrival
 62 and demographic stochasticity to jointly determine community assembly. One reason
 63 why the debate continues about which process dominates and under what conditions is
 64 that the relative importance of these contrasting processes has rarely (Cottenie 2005,
 65 Paine and Harms 2009) been quantified in a way that allows comparisons among studies
 66 (Condit et al. 2002, Clark and McLachlan 2003, Volkov et al. 2003, Gilbert and
 67 Lechowicz 2004, Harpole and Tilman 2006). To advance the debate we need a method
 68 of estimating the relative importance of these different processes that is comparable
 69 across locations, environments and spatial/temporal scales (Weiher et al. 2011). We
 70 develop such a method in this paper by extending the maxent model of trait-based
 71 community assembly presented in Shipley (2010b).

72 The model involves two spatial scales: the landscape containing a meta-
 73 community, whose species define the species pool, and one or more local communities

within this landscape. The meta-community is thus composed of local communities that can potentially exchange propagules (Leibold et al. 2004), and a species is a member of the species pool if it has a non-zero probability of dispersing from the meta-community into a local community. However, a species will still be missing from a local community if dispersal is a sufficiently rare event such that no immigrants have yet arrived, if the species is so poorly adapted to the local biotic and abiotic conditions that immigrants rarely reproduce and quickly die before being observed, or if demographic stochasticity results in local extinction.

Alternative model assumptions

First assume “local neutral” assembly, in which every species in the species pool has the same per-capita probabilities of dispersing from the meta-community to the local community and of surviving, growing and reproducing in the local community once present (Hubbell 2001). Given this assumption, species having more individuals in the meta-community will disperse more immigrants to the local community, even though the per capita probabilities of immigration are equal across species, and local dynamics would then proceed without regard to trait values or niche partitioning. This leads to the “neutral” prior expectation: that the expected relative abundance of each species in the local community is equal to its relative abundance in the meta-community. The actual relative abundance of each species in each local community will be a random walk around these expected values because of demographic stochasticity. As the spatial scale of the local community decreases, so also will the total number of individuals of various species in this local community. These smaller population sizes will increase the stochastic variation around the expected relative abundances of these species, and this

will reduce the observed correlation between local relative abundance and meta-community relative abundance. In “local neutral” assembly, since trait values are unrelated to local fitness differences, they will also be independent of local relative abundances; therefore the trait values of the average individual in the local community will be approximately equal to the average trait value in the meta-community. This is a restricted version of Hubbell’s (2001) neutral assumption because we make no assumptions about the causes of relative abundance at the landscape level, whereas Hubbell’s model assumes that neutral processes also determine the meta-community structure. In our “local neutral” model, the causes of relative abundance at the landscape level might be due to unanalyzed historical, phylogenetic or biogeographic factors but could also be due to niche-based processes. Thus, although dispersal limitation generates the similarity between the meta-community and local relative abundances in this “local neutral” model, it does not necessarily follow that dispersal limitation is independent of trait differences, only that trait differences among species co-occurring in the local community do not affect their local relative abundances. In other words, species’ traits provide no new information beyond that already encoded in the meta-community relative abundances.

Next, assume “local niche-based” assembly, in which the *per capita* probabilities of dispersal, survival, growth and reproduction of a species in a local community are completely determined by its functional traits in interaction with the local biotic and abiotic environment. If this were the case, then the relative abundance of the species in a local community would be entirely determined by the selective advantage of their traits, given the local environmental conditions. This follows because each species in the

species pool has a non-zero probability of dispersing to the local community and subsequent population dynamics of the species, once it has arrived, would be determined by trait-based demographic properties of growth, survival and reproduction. Once one knows the trait values of a species then knowledge of its relative abundance in the larger meta-community would be irrelevant. This does not preclude demographic stochasticity from affecting the abundance of species with small population sizes. Those species having adaptively advantageous traits would have more individuals, and so the trait value of the average individual in the local community will be biased towards the more locally well-adapted species.

Finally, assume “hybrid” local assembly: The abundance of a species in the meta-community affects its probability of arrival (i.e. dispersal limitation) in a local community because more abundant species in the meta-community will tend to produce more propagules; therefore meta-community abundances will contain relevant information. However, per capita probabilities of immigration will not be equal across species and these differences will be associated with trait adaptations relevant to dispersal. Subsequent local dynamics are also shaped by trait values and local environmental conditions. In this case, knowledge of both the meta-community relative abundance and the trait values of a species would be informative in predicting its relative abundance in the local community. Species that are abundant in the meta-community could have somewhat higher abundances in the local community than their traits would predict due to mass flow (Shmida and Wilson 1985) or because a competitively superior species may be absent from the local community due to dispersal limitation. Species that are rare in the meta-community could have somewhat lower abundances in the local

community than their traits would predict, because demographic stochasticity arising from a very small immigrant pool could force them to local extinction before their population sizes become large enough that trait selection dominates. The relative strength of dispersal limitation, trait-based filtering, and demographic stochasticity will determine the actual local community structure.

In this paper, we evaluate the relative importance of dispersal limitation, trait-based filtering, and demographic stochasticity in the assembly of species-rich tree communities of mature tropical rain forests in French Guiana for which extensive information is available on 17 functional traits. We further examine the extent to which the relative importance of each process is scale-dependent by conducting our analyses at three spatial scales.

METHODS

The data (Baraloto et al. 2010) come from a set of nine 1-ha tropical forest plots in French Guiana separated, on average, by 117 km (range: 0.5 to 300 km). Every tree of at least 10 cm in diameter at breast height was mapped and identified to species in 2007 and 2008. Because each individual was spatially referenced, we could also determine species relative abundances in subplots of 0.25 and 0.04 ha within each plot. These abundance data (number of individuals) were used to estimate relative abundances of each species both at the meta-community level (i.e. combining all nine plots) and at each of the three smaller spatial scales. Table 1 summarizes these data. There were 657 species in the meta-community, but only a small proportion of these were present in any single local community. Approximately 1500 tree species have been collected in all of

French Guiana, including central and southern regions that have a more distinctive flora (Sabatier and Molino, unpublished data), so our sample likely represents at least 50% of the regional pool. The percentage of these 657 species that were missing from a single local community at the three spatial scales was 77% (1 ha), 89% (0.25 ha) and 97% (0.04 ha).

We measured values of 17 traits (Table 2) on almost every individual tree and calculated the average value per trait for each of the 657 species in the species pool, as detailed in Baraloto et al (2010). This was done in two ways. First, we calculated the mean values per species averaged over all measured individuals per species over all nine plots (“meta-community trait means”). Second, we calculated the mean values per species averaged over all measured individuals per species in each 1-ha plot (“plot-specific trait means”); when a species was absent from a given plot its meta-community trait means were used. Although leaf traits were measured on every individual, wood traits and chemistry were not. We therefore estimated unobserved trait values using Multiple Imputation with Chained Equations (MICE), as implemented in the mice package of R (Team 2008). Unobserved values were estimated through predictive mean matching using all other data as predictors. Community-weighted trait means,

$\bar{t}_{jk} = \sum_{i=1}^S r_{ik} t_{ij}$, were calculated for each of the j traits in each of the k local communities (i.e. plots or subplots) based on each of the i species in the species pool; r_{ik} is the observed relative abundance of species i in local community k ; for plot-level trait means we used t_{ijk} , i.e. trait j of species i in plot k instead of t_{ij} . These average trait values were used as constraints in the maxent model.

Predicted relative abundances from the maxent model were obtained using the maxent function of R's FD library (Laliberté and Shipley 2009). These predicted relative abundances are simultaneously maximum entropy estimates and maximum likelihood estimates; this is explained more fully in the electronic appendix. Null probabilities for the statistical tests were obtained using the permutation test of Shipley (2010c) as implemented by the maxent.test function of the FD library. Because of the numerically intensive nature of this test when applied to such a large data set we fixed a significance level of 0.05 and stopped the permutation runs when the 95% confidence intervals of the estimated null probability were outside the 0.05 level (Besag and Clifford 1991). For ease of comparison, and because the subplots smaller than 1 ha were clearly spatially dependent, we calculated the null probabilities separately for each of the nine plots even when the models were fit separately for each subplot, as described in Shipley (2010c). We used a maximally uninformative (uniform) prior when testing the pure local trait-based model and used a local neutral prior, consisting of the observed relative abundances at the meta-community level, for the hybrid model. These priors are explained below.

Decomposing the information content of a relative abundance distribution to infer the contribution of traits, dispersal limitation, and other factors.

We estimate the relative importance of stochastic and niche-based processes by modeling the pure local neutral, pure local trait-based, and the hybrid assumptions of community assembly using the maximum entropy CATS model (Community Assembly through Trait Selection) introduced in Shipley et al. (2006) and developed in detail in Shipley (2010b). Our approach to partitioning the total information content into

components implied by these three models is similar to Borcard *et al.*'s (1992) method of partitioning the variance in constrained ordinations of abundance data into purely spatial, purely environmental, jointly spatial and environmental, and undetermined components.

The supplement in Ecological Archives summarizes the model.

The CATS model has three inputs: the trait matrix, the prior distribution and the community-weighted trait means (trait constraints). The trait matrix lists the trait values of each species in the species pool; these values are species' means estimated over either the full meta-community or separately in each local community. The prior distribution is either the neutral prior (if non-trait based dispersal limitation is assumed to occur) or a maximally uninformative uniform prior (if non-trait based dispersal limitation is assumed not to occur). The trait constraints for a given local community are a vector of observed community-weighted trait means, which give the trait values of an average individual in the local community. The model finds the predicted relative abundances of each species in the species pool in each local community such that these predictions are consistent with the community-weighted trait means in the local community while simultaneously having maximum relative entropy with respect to (thus as close as possible to) the given prior.

As such, the CATS model contrasts two opposing forces and allows one to estimate the relative importance of each. As the importance of local trait-based filtering increases, the proportion of the information contained in observed relative abundances that is accounted for by the trait model ("predictive ability") increases relative to the predictive ability given only the prior (either a uniform or neutral prior). As the importance of dispersal limitation from the landscape increases, the predictive ability of

the model given only the neutral prior increases. Given a pure niche-based process of community assembly, the model will have maximal predictive ability (assuming that all relevant traits are incorporated) whereas the neutral prior alone will have no predictive ability. A strictly local neutral process of community assembly results in a model with no predictive ability beyond that provided solely by the neutral prior; in this scenario the predictive ability of the neutral prior will decrease as the importance of demographic stochasticity increases since the latter increases random variation around the expected neutral relative abundance distribution. Thus, the predictive ability of the model can be used to estimate the relative importance of trait-based and stochastic processes occurring during community assembly, and this predictive ability can be compared to alternative null distributions that incorporate different assumptions about the importance of dispersal limitation.

The proportion of the uncertainty in observed relative abundances that is accounted for by the model is quantified by Cameron and Windmeijer's (1997) generalization (R_{KL}^2 , equation 1), based on the Kullback-Leibler divergence, of the classic R^2 measure of goodness of fit between observed and model predicted values that is applicable in linear regression models with a normal error structure. This R_{KL}^2 generalization has all of the same properties possessed by the classic R^2 and so R_{KL}^2 quantifies the proportion of uncertainty explained by the type of fitted exponential regression model that is implied by the CATS maximum entropy model. The supplement justifies this measure in detail. The total species pool is S (here 657), the total number of local communities is C whose value depends on the spatial scale, o_{ij} is the observed relative abundance of species i in local community j , p_{ij} is the model prediction of the

relative abundance given the different assumptions of community assembly, and $q_{i,0}=1/S$ is the predicted relative abundance given only the maximally uninformative prior.

$$R_{KL}^2 = 1 - \frac{\sum_{j=1}^C \sum_{i=1}^S o_{ij} \ln \left(\frac{o_{ij}}{p_{ij}} \right)}{\sum_{j=1}^C \sum_{i=1}^S o_{ij} \ln \left(\frac{o_{ij}}{q_{i,0}} \right)}$$

Equation 1

We decompose the total explained uncertainty in the observed relative abundances into separate components using the estimates of R_{KL}^2 given four different models, while correcting for model bias (equations 2 – 5). The supplement develops and justifies the equations for this decomposition; only the final equations are presented here.

The data are fit to four versions of the full CATS model. The first model, estimating only model bias and giving $\hat{R}_{KL,null}^2 (uniform)$, is fit using a maximally uninformative (i.e. uniform) prior and permuted trait values. This removes any association between traits and relative abundances and excludes any contribution due to dispersal limitation (because a uniform prior assumes equal contributions of propagules of all species in the species pool) while maintaining any contribution to R_{KL}^2 due to model structure. The second model, estimating only dispersal limitation due to processes occurring at the landscape level and giving $\hat{R}_{KL,null}^2 (neutral)$, is fit using a neutral prior and randomly permuted trait values. This again removes any association between traits and relative abundances but includes any contribution due to non-trait related dispersal limitation. The third model, estimating only trait filtering and giving $R_{KL}^2 (uniform, traits)$, is fit using the maximally uninformative prior and non-permuted trait values. This incorporates any association between traits and relative abundances but excludes any

contribution due to non-trait related dispersal limitation. The final model, estimating both dispersal limitation and trait filtering, is fit using the neutral prior and non-permuted trait values and gives $R_{KL}^2(neutral, traits)$.

Equations 2 to 5 decompose the total explained uncertainty into five components:

(i) model bias ($\hat{R}_{KL,null}^2(uniform)$), (ii) the association due uniquely to trait-based filtering, (iii) the association due uniquely to meta-community processes (non-trait based dispersal), (iv) the association due jointly to meta-community processes and local trait-based filtering and (v) the remaining unexplained variance that can be provisionally ascribed to demographic stochasticity.

Proportion of biologically relevant uncertainty explained uniquely by local trait

$$\text{constraints: } \frac{R_{KL}^2(neutral, traits) - \hat{R}_{KL,null}^2(neutral)}{1 - \hat{R}_{KL,null}^2(uniform)} \quad \text{Equation 2}$$

Proportion of biologically relevant uncertainty explained uniquely by the neutral prior:

$$\frac{R_{KL}^2(neutral, traits) - R_{KL}^2(uniform, traits)}{1 - \hat{R}_{KL,null}^2(uniform)} \quad \text{Equation 3}$$

Proportion of biologically relevant uncertainty explained jointly by both the trait

constraints and the neutral prior:

$$\frac{\hat{R}_{KL,null}^2(neutral) + R_{KL}^2(uniform, traits) - R_{KL}^2(neutral, traits) - \hat{R}_{KL,null}^2(uniform)}{1 - \hat{R}_{KL,null}^2(uniform)}$$

Equation 4

Biologically relevant unexplained information:

$$\frac{1 - (R_{KL}^2(neutral, traits) - \hat{R}_{KL,null}^2(uniform)) - \hat{R}_{KL,null}^2(uniform)}{1 - \hat{R}_{KL,null}^2(uniform)} = \frac{1 - R_{KL}^2(neutral, traits)}{1 - \hat{R}_{KL,null}^2(uniform)}$$

Equation 5

RESULTS

Trait-based filtering explained 72% of the information contained in relative abundances at the 1-ha scale, and this percentage decreased to 49% at the 0.04-ha scale (Table 3). The hybrid model, which takes both dispersal limitation and functional traits into account, explained 82% of the information in relative abundances at the 1-ha scale, decreasing to 55% at the 0.04ha scale. The predictive ability of both the trait-only and the hybrid models was always greater using plot-specific trait means than meta-community trait means. Dispersal limitation, as inferred from the neutral prior, explained at least 39% of the information at all scales, but decreased slightly in importance as the spatial scale decreased.

Ecological inferences from the models are facilitated by considering only biologically relevant information on relative abundance. There was an association between the meta-community relative abundances (the neutral prior) and the site-specific trait constraints such that, after removing model bias, only 34% of the information at the 1-ha scale was uniquely attributable to niche-based processes. A similar amount of information (36%) was due jointly to traits and the neutral prior, and the two cannot be separated (Figure 1). Also because of the meta-community abundance – local trait constraint association, the proportion of the information explained uniquely by the neutral prior was only 10% at the 1-ha scale and remained approximately constant at all spatial scales. The proportion of the total biologically relevant information that remained unexplained, and was tentatively ascribed to demographic stochasticity, was only 20% at the 1-ha scale but increased to 50% at the 0.04-ha scale (Figure 1).

DISCUSSION

The method developed here estimates the relative importance of local trait-based filtering, dispersal limitation and demographic stochasticity to community assembly. The partitioning of information reflects the contribution of each process in determining the relative abundance of each species in the species pool within a local community. The minimal data requirements are measures of relative abundance in the local community and trait values of each species in the meta-community species pool. We applied this method to evaluate the degree to which the relative importance of these processes changes with spatial scale. Our method would also facilitate evaluating the changes in relative importance of these processes along environmental gradients. Importantly, the relationship between these population processes and the resulting statistical model is explicit, because the underlying model can be mathematically linked to population dynamics and quantitative selection on traits (Shipley 2010a, chapter 5, 2010b).

Gilbert and Lechowicz (2004) and Cottenie (2005) , based on Borcard et al. (1992), proposed a related method of decomposing species relative abundances that can also be linked to niche-based and neutral processes, but which does not use functional traits. This method requires information on (i) species abundance across many local communities (the community matrix \mathbf{C}), (ii) explicitly measured environmental variables in each local community (the environmental matrix \mathbf{E}) and (iii) the spatial configuration of the local communities (the spatial matrix \mathbf{S}). Cottenie's variance decomposition consists of a (partial) redundancy analysis of \mathbf{C} given either (i) \mathbf{E} , (ii) \mathbf{S} , (iii) $\mathbf{E} + \mathbf{S}$, (iv) \mathbf{E} controlling for \mathbf{S} ($\mathbf{E}|\mathbf{S}$) or (v) \mathbf{S} controlling for \mathbf{E} ($\mathbf{S}|\mathbf{E}$). The assumption is that the partial redundancy analysis of \mathbf{C} given the environmental matrix after controlling for the spatial

matrix (**E|S**) reflects niche-based species filtering along the gradient (i.e. β -diversity) while the partial redundancy analysis of C given the spatial matrix after controlling for the environmental matrix (**S|E**) reflects neutral processes or patch dynamics but not niche-based processes. The E+S fraction, which is often the largest one, cannot be used to make conclusive statements about the relative importance of either E or S. However, unlike the CATS model upon which our method is based (Shipley 2010b, chapter 5), the statistical model has not been explicitly linked to these underlying population processes and so one cannot know to what degree the assumptions are reasonable or robust. We could not apply Cottenie's method to our data because we could not identify clear abiotic gradients that differentiate our sites and that clearly drive the underlying dynamics. There are therefore three differences between our method and that proposed by Cottenie: (i) the type of data required, (ii) the link between the statistical patterns and the demographic processes and (iii) the emphasis on explaining differences in relative abundance within a local community (us) versus explaining differences between communities (Cottenie). Cottenie's method and ours should therefore be viewed as partly complementary.

Our observational method contrasts with experimental approaches, including that of adding propagules of new species to a local community. The survival and reproduction of experimental colonists in a local community can provide evidence of dispersal limitation, but the method cannot quantify the relative importance of niche-based and neutral processes. By necessity, most such experimental studies have focused on taxa where dispersal is easily manipulated, where the organisms complete their life cycle rapidly, and where community dynamics are very rapid, as shown by the meta-

analysis of Myers and Harms (2009). Although the link between pattern and process is more direct using this experimental approach, it is difficult to apply to many natural systems and cannot provide a quantitative estimate of the relative importance of trait-based filtering and dispersal limitation. Furthermore, because such experiments are of relatively short duration, they can miss longer-term processes of trait-based filtering (Paine and Harms 2009).

In previous applications of the CATS model to a diverse set of herbaceous plant communities, a strong signal of trait-based community assembly was detected, which suggested a dominant role for trait-based filtering (Shipley et al. 2006, Sonnier 2009, Mokany and Roxburgh 2010, Radovski 2010, Sonnier et al. 2010, Merow et al. 2011, Shipley et al. 2011). None of these studies, however, decomposed the information in species relative abundances and explicitly estimated the information explained by dispersal limitation. In the current study, contrastingly, trait-based filtering explained a maximum of only 34% of the information, and its contribution was less at smaller scales. The evidence of a predominance of trait-based processes provided by these previous studies is therefore still tentative, especially because such a sample of communities is much too small and unrepresentative to draw any general conclusions. In particular, community assembly in tropical forests seems to be particularly sensitive to stochastic processes; indeed, these are the communities that inspired Hubbell's (2001) neutral model. The large number of species in these forests that coexist at small spatial scales, without large differences in functional traits, and without pronounced environmental gradients, suggest that niche-based assembly processes will be weak. On the other hand, there is also empirical evidence of trait-based differences in performance and distribution

among tropical tree species (Kraft et al. 2008, Poorter et al. 2008, Kraft and Ackerly 2010) that point to niche-based processes. What then is the relative importance of trait-based species' filtering vs. dispersal limitation vs. demographic stochasticity in structuring these tropical forests?

Our results help to answer this question. Local community structure in the studied tropical rain forests is consistent with neither a purely neutral nor a purely niche-based process. There was clear statistical evidence of trait-based filtering in the local communities, consistent with the earlier results of Paine et al. (2011) who studied the same forests. On the other hand, only 34% of the information in relative abundances was associated uniquely with trait-based filtering at the 1-ha scale, and this percentage decreased with decreasing spatial scale. Dispersal limitation was approximately constant across spatial scales but was always less important than trait-based filtering based on the more informative plot-specific trait means. Because there was redundant information contained in the neutral prior and the site-level trait constraints, this proportion of the biologically relevant information cannot be unambiguously ascribed to only local niche-based or to dispersal limitation alone. However, the fact that the pure dispersal signal remained essentially constant as the spatial scale of the local community increased, whereas the pure trait-based signal increased, suggests that the redundant information is due to species' traits affecting (or being correlated with) other landscape factors that affect relative abundances in the meta-community. It is impossible to identify these landscape factors given our data, but one possibility is that the relative abundance of the species in the meta-community partly reflects the relative abundance of the different environmental conditions in the landscape to which these species are adapted. This

would explain both why there was redundant information between site-specific traits and meta-community abundances and why the pure dispersal limitation proportion of the information did not change much at the different spatial scales. However, it is also possible that the meta-community relative abundances simply reflect biogeographic processes of speciation (Hubbell 2001) and that our measured traits partly track this phylogenetic signal.

The unexplained variation, which we provisionally ascribe to demographic stochasticity, showed the opposite trend. At the 1-ha scale, with over 500 trees on average per local community, this unexplained variation was slightly less than that due only to dispersal limitation; but at the smallest spatial scale, with only 30 trees on average, it was three times more important. This is what one would expect if this unexplained variation was primarily due to demographic stochasticity: as population sizes decrease, chance fluctuations in individual birth and death rates would increasingly dominate population dynamics even if the species mean probabilities of reproduction, survival and growth were influenced by traits. This scale-dependence of the relative importance of neutral and niche processes has already been reported (Karst et al. 2005, Laliberté et al. 2009). However, the interpretation of the biologically unexplained information with demographic stochasticity must be made with care because if there are important unmeasured traits that were not strongly correlated with our measured traits then this would also contribute to the biologically unexplained information. Increasing the number of traits in the model will not artificially inflate the explained information because this source of bias is corrected using $\hat{R}_{KL,null}^2 (uniform)$. Sonnier et al. (2011) describe how to select a parsimonious set of traits using the CATS model. Similarly, the

importance of demographic stochasticity will be inflated if the values used in the neutral prior are poor estimates of the relative abundances of potential immigrants in the meta-community. Our estimate should therefore be considered as an upper bound on the importance of demographic stochasticity.

There are at least three sources of error that should be considered in our analysis. First, there might be strong but unknown environmental gradients that would result in different selection pressures on the traits within the same local community and that largely cancel out when averaged over the spatial scale of the entire local community. We have not found any such environmental gradients within the local communities and doubt that such strong gradients exist. Since the length of environmental gradients is often related to the spatial scale over which the environment is sampled (Bell and Lechowicz 1994) then a local community defined at a larger spatial scale would be more likely to contain a larger environmental gradient within it. If so then one would expect an increase in the strength of trait-based filtering at smaller spatial scales since any environmental gradients at such small scales would be less pronounced; this expectation is contrary to our results.

Second, we might have included many species in our species pool that have no possibility (as opposed to simply a low probability) of reaching our local communities. If the true species pool is greatly overestimated then this necessarily leads to prediction errors in the maxent model (Sonnier et al. 2010). However, all but 30 species in our pool occur across the general region, all are likely capable of growing in all plots in the absence of competition, and there are no obvious non-biological barriers that would make it impossible for any species to immigrate to each of the local communities. Properly

identifying all those species that could reasonably disperse into, and survive, the abiotic conditions of the local communities (i.e. the species pool) is a difficult and unsolved problem. This is a weakness of the CATS model as currently implemented.

Finally, there might be strong trait-based filtering on traits not measured by us and that are largely independent of those that we did measure. The 17 measured functional traits expand on a well-accepted leaf economics spectrum (Wright et al. 2004) to define tropical tree tissue strategies ((Baraloto et al. 2010). Note that some of the previously cited studies using the CATS model detected much stronger associations between traits and relative abundance with as few as 4 or 5. It is possible (even likely) that our analysis somewhat underestimates the degree of trait-based filtering; for instance, we do not have traits related to herbivore or pathogen defense or below-ground traits, and these might be uncorrelated with our measured traits.

Despite the limitations discussed here, our method provides estimates of the relative importance of traits, dispersal limitation and demographic stochasticity to the assembly of the tree communities in these species-rich forests. Our method should allow researchers to quantitatively compare across studies, locations, landscape features and spatial/temporal scales. In this way we can move beyond the neutral/niche debate to study how such properties combine to determine the relative importance of niche-based and neutral processes during community assembly in different environmental, spatial and temporal contexts.

ACKNOWLEDGEMENTS

481 This research was funded by the Natural Sciences and Engineering Research
 482 Council of Canada (NSERC), the Agence National de la Recherche (ANR) of France, the
 483 US National Science Foundation (DEB 0743103) and an INRA Package grant to CB. We
 484 thank all participants of the project BRIDGE who participated in field and laboratory
 485 collection and treatment of specimens. Field research was facilitated by the Guyafor
 486 permanent plot network in French Guiana which is managed by CIRAD, ONF and
 487 CNRS. We thank Bruno Hérault, Eric Marcon, Vivien Rossi, Evan Weiher and one
 488 anonymous reviewer for valuable comments.

489

preprint

490

491

References

492

493

494

495

496

497 Baraloto, C., C. E. T. Paine, S. Patino, D. Bonal, B. Hérault, and J. Chave. 2010.

498

499

Functional trait variation and sampling strategies in species-rich plant communities. *Functional Ecology* **24**:208-216.

500

501

Bell, G. 2000. The distribution of abundance in neutral communities. *Am Nat* **155**:606-617.

502

503

504

505

Bell, G. and M. J. Lechowicz. 1994. The flip side: manifestations of how plants perceive patchiness at different scales. Pages 391-414 *in* M. M. Caldwell and R. W. Pearcy, editors. *Exploitation of environmental heterogeneity by plants. Ecophysiological processes above and below ground*. Academic Press, NY.

506

507

Besag, J. and P. Clifford. 1991. Sequential Monte Carlo p-values. *Biometrika* **78**:301-304.

508

509

Borcard, D., P. Legendre, and P. Drapeau. 1992. Partialling out the spatial component of ecological variation. *Ecology* **73**:1045-1055.

510

511

512

Cameron, C. A. and F. A. G. Windmeijer. 1997. An R-squared measure of goodness of fit for some common nonlinear regression models. *Journal of Econometrics* **77**:329-342.

513

514

Clark, J. S. and J. S. McLachlan. 2003. Stability of forest biodiversity. *Nature* **423**:635-638.

- 515 Condit, R., N. Pitman, E. G. Leigh Jr, J. Chave, J. Terborgh, R. B. Foster, P. V. Núñez, S.
516 Aguilar, R. Valencia, G. Villa, H. C. Muller-Landau, E. Losos, and S. P. Hubbell.
517 2002. Beta-diversity in tropical forest trees. *Science* **295**:666-669.
- 518 Cottenie, K. 2005. Integrating environmental and spatial processes in ecological
519 community dynamics. *Ecology Letters* **8**:1175-1182.
- 520 Gilbert, B. and M. J. Lechowicz. 2004. Neutrality, niches, and dispersal in a temperate
521 forest understory. *Proceedings of the National Academy of Sciences of the United*
522 *States of America* **101**:7651-7656.
- 523 Gravel, D., C. Charles D, M. Beaudet, and C. Messier. 2006. Reconciling niche and
524 neutrality: the continuum hypothesis. Pages 399-409. Wiley-Blackwell.
- 525 Harpole, W. S. and D. Tilman. 2006. Non-neutral patterns of species abundance in
526 grassland communities. *Ecol Lett* **9**:15-23.
- 527 Hubbell, S. P. 2001. The unified neutral theory of biodiversity and biogeography.
528 Princeton University Press.
- 529 Karst, J., B. Gilbert, and M. J. Lechowicz. 2005. Fern community assembly: The roles of
530 chance and the environment at local and intermediate scales. *Ecology* **86**:2473-
531 2486.
- 532 Kraft, N. J. B. and D. D. Ackerly. 2010. Functional trait and phylogenetic tests of
533 community assembly across spatial scales in an Amazonian forest. *Ecological*
534 *Monographs* **80**:401-422.
- 535 Kraft, N. J. B., R. Valencia, and D. D. Ackerly. 2008. Functional traits and niche-based
536 tree community assembly in an Amazonian forest. *Science* **322**:580-582.

- 537 Laliberté, E., A. Paquette, P. Legendre, and A. Bouchard. 2009. Assessing the scale-
538 specific importance of niches and other spatial processes on beta diversity: A case
539 study from a temperate forest. *Oecologia* **159**:377-388.
- 540 Laliberté, E. and B. Shipley. 2009. FD: Measuring functional diversity from multiple
541 traits, and other tools for functional ecology. R package. Vienna, Austria.
- 542 Leibold, M. A., M. Holyoak, N. Mouquet, P. Amarasekare, J. M. Chase, M. F. Hoopes,
543 R. D. Holt, J. B. Shurin, R. Law, D. Tilman, M. Loreau, and A. Gonzalez. 2004.
544 The metacommunity concept: a framework for multi-scale community ecology.
545 *Ecology Letters* **7**:601-613.
- 546 Lotka, A. J. 1925. Elements of physical biology. Williams & Williams, Baltimore.
- 547 MacArthur, R. H. 1972. Geographical Ecology: Patterns in the Distribution of Species.
548 Harper and Row, N.Y.
- 549 Merow, C., A. Latimer, and J. Silander, John. 2011. Can entropy maximization use
550 functional traits to explain species abundances? A comprehensive evaluation.
551 *Ecology* **92**:1523-1537.
- 552 Mokany, K. and S. H. Roxburgh. 2010. The importance of spatial scale for trait-
553 abundance relations. *Oikos* **119**:1504-1514.
- 554 Myers, J. A. and K. E. Harms. 2009. Seed arrival, ecological filters, and plant species
555 richness: A meta-analysis. *Ecology Letters* **12**:1250-1260.
- 556 Paine, C. E. T., C. Baraloto, J. Chave, and B. Hérault. 2011. Functional traits of
557 individual trees reveal ecological constraints on community assembly in tropical
558 rain forests. *Oikos* **120**:720-727.

- 559 Paine, C. E. T. and K. E. Harms. 2009. Quantifying the effects of seed arrival and
560 environmental conditions on tropical seedling community structure. *Oecologia*
561 **160**:139-150.
- 562 Poorter, L., S. J. Wright, H. Paz, D. D. Ackerly, R. Condit, G. Ibarra-Manríquez, K. E.
563 Harms, J. C. Licona, M. Martínez-Ramos, S. J. Mazer, H. C. Muller-Landau, M.
564 Peña-Claros, C. O. Webb, and I. J. Wright. 2008. Are functional traits good
565 predictors of demographic rates? Evidence from five neotropical forests. *Ecology*
566 **89**:1908-1920.
- 567 Radovski, M. 2010. Testing a trait-based model of fern community assembly. McGill
568 University, Montreal, Canada.
- 569 Shipley, B. 2010a. Community assembly, natural selection and maximum entropy
570 models. *Oikos* **119**:604-609.
- 571 Shipley, B. 2010b. From plant traits to vegetation structure : chance and selection in the
572 assembly of ecological communities. Cambridge University Press, Cambridge.
- 573 Shipley, B. 2010c. Inferential permutation tests for maximum entropy models in ecology.
574 *Ecology* **91**:2794-2805.
- 575 Shipley, B., D. C. Laughlin, G. Sonnier, and R. Otfinowski. 2011. A strong test of the
576 maximum entropy model of trait-based community assembly. *Ecology* **92**:507-
577 517.
- 578 Shipley, B., D. Vile, and É. Garnier. 2006. From plant traits to plant communities: A
579 statistical mechanistic approach to biodiversity. *Science* **314**:812-814.
- 580 Shmida, A. and M. V. Wilson. 1985. Biological determinants of species diversity. *Journal*
581 *of Biogeography* **12**:1-20.

- 582 Sonnier, G. 2009. Des traits des espèces à la structure des communautés végétales, vers
583 une approche fonctionnelle et prédictive de l'écologie des communautés.
584 Université de Sherbrooke Editor, Sherbrooke.
- 585 Sonnier, G., B. Shipley, A. Fayolle, and M. L. Navas. 2011. Quantifying trait selection
586 driving community assembly: A test in herbaceous plant communities under
587 contrasted land use regimes. *Oikos* **In press**.
- 588 Sonnier, G., B. Shipley, and M. L. Navas. 2010. Plant traits, species pools and the
589 prediction of relative abundance in plant communities: a maximum entropy
590 approach. *J Veg Sci* *J Veg Sci* **21**:318-331.
- 591 Tansley, A. G. 1920. The classification of vegetation and the concept of development.
592 *Journal of Ecology* **8**:118-149.
- 593 Team, R. D. C. 2008. R: A language and environment for statistical computing. R
594 Foundation for Statistical Computing, Vienna.
- 595 Tilman, D. 1982. Resource competition and community structure. Princeton University
596 Press, Princeton.
- 597 Tilman, D. 1988. Plant strategies and the dynamics and structure of plant communities.
598 Princeton University Press, Princeton.
- 599 Volkov, I., J. R. Banavar, S. P. Hubbell, and A. Maritan. 2003. Neutral theory and
600 relative species abundance in ecology. *Nature* **424**:1035-1037.
- 601 Volterra, V. 1926. Variazioni e fluttuazioni del numero d'individui in specie animali
602 conviventi. *Mem.R.Acad.dei Lincei (Ser.6)* **2**:31-113.
- 603 Volterra, V. 1931. Leçons sur la Théorie Mathématique de la Lutte pour la Vie. Gauthier-
604 Villars, Paris.

605 Weiher, E., D. Freund, T. Bunton, A. Stefanski, T. Lee, and S. Bentivenga. 2011.
 606 Advances, challenges and a developing synthesis of ecological community
 607 assembly theory. *Philosophical Transactions of the Royal Society B: Biological*
 608 *Sciences* **366**:2403-2413.

609 Wright, I. J., P. B. Reich, M. Westoby, D. D. Ackerly, Z. Baruch, F. Bongers, J.
 610 Cavender-Bares, T. Chapin, J. H. C. Cornelissen, M. Diemer, J. Flexas, E.
 611 Garnier, P. K. Groom, J. Gulias, K. Hikosaka, B. B. Lamont, T. Lee, W. Lee, C.
 612 Lusk, J. J. Midgley, M. L. Navas, Ü. Niinemets, J. Oleksyn, H. Osada, H. Poorter,
 613 P. Pool, L. Prior, V. I. Pyankov, C. Roumet, S. C. Thomas, M. G. Tjoelker, E. J.
 614 Veneklaas, and R. Villar. 2004. The worldwide leaf economics spectrum. *Nature*
 615 **428**:821-827.

616
 617
 618
 619

Table 1. Summary of the data used to fit the various models listing the number of separate local communities fit by the models at each spatial scale as well as the median (minimum - maximum) total number of individuals and trees species per community.

Level	Number of local communities	Median number of individuals	Median number of species
Meta-community	1	4557	657
1-ha	9	507 (406 - 632)	150 (145 - 202)
0.25 ha	36	123 (64 - 160)	69 (42 - 94)
0.04 ha	224	29 (5 - 32)	17 (4 - 28)

Table 2. Functional traits measured on each of 657 species, along with the units, number of individuals measured for traits, and the median and range of observed species-mean trait values.

Functional trait	unit	N	Median	Range in dataset
Foliar $\delta^{13}\text{C}$ composition	‰	2946	-32.070	-36.44 – -25.67
Foliar C:N ratio	g g^{-1}	2947	24.01	7.79 – 59.77
Foliar K concentration	%	933	0.05	0.114 – 2.235
Foliar N concentration	%	2948	02.0	0.762 – 6.190
Foliar P concentration	%	933	0.1	0.024 – 0.251
Laminar chlorophyll content	$\mu\text{g mm}^{-2}$	4611	70.5	10.3 – 255.1
Laminar total surface area	cm^2	4587	72.1	2.032 – 643700
Laminar toughness	N	4590	1.650	0.22 – 13.06
Leaf tissue density	g cm^{-3}	4540	0.042	0.008 – 0.287
Leaflet surface area	cm^2	4587	52.23	0.018 – 3218
NH_4^+ utilization	%	4587	47	1 – 100
Specific Leaf Area (SLA)	$\text{cm}^2 \text{g}^{-1}$	4577	10.68	1.77 – 47.41
Trunk bark thickness	mm	3805	4	0.0 – 53.0
Trunk xylem density	g cm^{-3}	2844	0.65	0.23 – 0.98
Trunk xylem moisture content	%	2256	61.4	17 – 287
Twig bark thickness	mm	2369	1.95	0.07 – 7.31
Twig xylem density	g cm^{-3}	2390	0.616	0.19 – 0.96

630 Table 3. Estimated proportions (R_{KL}^2) of the total information in the local observed
631 relative abundances in 657 tree species in 9 local sites in French Guiana that is accounted
632 for by various maxent models. Shown are the results when the mean trait value of each
633 species is calculated over all 9 sites (meta-community trait means) and when the mean is
634 calculated separately for each site (site-level trait means). Values in parentheses give the
635 number of sites out of 9 for which the trait constraints significantly improved the model
636 over the prior alone.

R_{KL}^2 value	1-ha scale	0.25-ha scale	0.04-ha scale
Site-level trait means			
$\hat{R}_{KL,null}^2$ (<i>uniform</i>)	0.066	0.081	0.162
Model bias			
$\hat{R}_{KL,null}^2$ (<i>neutral</i>)	0.497	0.410	0.385
Pure neutral			
R_{KL}^2 (<i>uniform, traits</i>)	0.721 (9/9)	0.580 (9/9)	0.487 (9/9)
Pure trait selection			
R_{KL}^2 (<i>neutral, traits</i>)	0.815 (9/9)	0.658 (9/9)	0.554 (9/9)
Hybrid model			
Meta-community trait means			
$\hat{R}_{KL,null}^2$ (<i>uniform</i>)	0.056	0.073	0.151
Model bias			
$\hat{R}_{KL,null}^2$ (<i>neutral</i>)	0.501	0.406	0.387

Pure neutral			
R_{KL}^2 (<i>uniform, traits</i>)	0.146 (4/9)	0.134 (4/9)	0.179 (3/9)
Pure trait selection			
R_{KL}^2 (<i>neutral, traits</i>)	0.596 (7/9)	0.485 (7/9)	0.429 (6/9)
Hybrid model			

637

638

preprint

639 Figure captions

640 Figure 1. Estimated proportions of the biologically relevant information in the local
 641 observed relative abundances in 657 tree species in 9 local sites in French Guiana that is
 642 accounted for by various maxent models. The information is explained by local trait-
 643 based filtering, dispersal limitation from the meta-community, unexplained information
 644 attributed to demographic stochasticity and joint trait – dispersal effects. These are
 645 plotted at three spatial scales and based on species’ trait means calculated separately for
 646 each local community (“plot-specific trait means”) and trait means over the entire meta-
 647 community.

648

preprint

Figure 1

